DOI：10.13998/j.cnki.issn1002-1248.23-0303

基于响应面分析的国家预印本平台自组织同行评议参数设计

王凌峰，王申鹏
（桂林电子科技大学 商学院，桂林 541004）

摘　要：[目的/意义]为保障国家预印本平台论文质量，需引入同行评议。近年来自组织同行评议逐渐受到关注。为达到国家预印本平台管理机构预期的性能表现，需要研究自组织同行评议关键参数优化设计。[方法/过程]以论文评议完成率（CR）、评议人任务分配均衡度（TBD）、论文平均评议时间（A_r_time）为自组织同行评议性能评价的3个因变量，评议人休息时间、评议人拒审率、作者人数、评议资格比例4个关键参数为自变量，对Box-Behnken抽样设计得到数据进行响应面分析，得到性能评价因变量与关键参数自变量的量化表达关系式。[结果/结论]通过4个关键参数自变量取值不同组合，国家预印本论文平台可以调整自组织同行评议的3个性能参数表现，保障国家预印本平台发布论文平均质量。

关键词：国家预印本平台；论文质量；自组织同行评议；Box-Behnken 设计；响应面分析
中图分类号：G35　　文献标识码：A　　文章编号：1002-1248（2023）07-0075-10

引用本文：王凌峰，王申鹏. 基于响应面分析的国家预印本平台自组织同行评议参数设计[J]. 农业图书情报学报，2023，35(7)：75-84.

1 引　言

科技信息是国家创新体系关键支撑要素，2020年中国颁发的《关于制定国民经济和社会发展第十四个五年规划和二〇三五年远景目标的建议》提出建设国家科研论文和科技信息高端交流平台，并将其列为“强化国家战略科技力量”七要点之一，这是中国首次将科技信息纳入中央级文件。作为落实文件上述要求的具体体现，科技部直属中国科学技术信息研究所建设了国家科研论文集中发布平台（China Open Access Archive，CoaA），该平台互联网址为 https://coaa.istic.ac.cn/，该平台已于2022年底上线。国家科研论文集中发布平台目前包括3个子平台——国家预印本平台、国家科技期刊平台、国家知识仓储平台。其中，国家预印本平台互联网为 https://coaa.istic.ac.cn/preprint。

国家预印本平台旨在推进中国科研论文快速发布和开放共享、实现原创科研论文“中国首发、全球知

收稿日期：2023-05-16
基金项目：国家自然科学基金项目“预印本数据库环境下开放同行评议机制研究”（71563007）；桂林电子科技大学研究生教育创新计划基金项目“预印本 2.0 与期刊评议速度差异定量建模分析”（2022YCXS120）
作者简介：王凌峰（1976-），男，博士，研究员，硕士生导师，研究方向为科技管理与创新。王申鹏（1997-），男，硕士研究生，研究方向为科技管理与创新

DOI：10.13998/j.cnki.issn1002-1248.23-0303

晓”。国家预印本平台接收中英文预印本科研论文，作者实名注册后即可投稿，文章通过最必要的基本审核后可获得 DOI 编号。为保障作者首发权，平台向首发论文提供电子收录证书。国家预印本平台同时整合中国目前最有代表性的多个预印本网站——教育部科技发展中心主办的中国科技论文在线（目前论文数量最多)、中国科学院文献情报中心主办的中国科学院预印本、中国科技信息研究所主办的中国预印本服务系统(该系统官网通告 2023 年 4 月 1 日停止接收稿件)。

国家预印本平台的高端定位要求其具有较完备的论文质量保障机制[1]。但是，预印本缺乏期刊的同行评议机制，故到目前为止，世界主要科研国家的科研人事管理部门在职称晋升、课题申报等诸多研究者关心的个人利益场合并没有将预印本文章与期刊文章等同视之。个人利益关切驱动下，大多数研究者虽然态度上支持预印本，但实际上并不向预印本提交文章。因此，尽管欧美代表性学科性预印本数据库（如数学、物理、计算机领域 arXiv，生物学领域 bioRiv）目前每年提交文章数量数以万计，但在数学、物理、计算机、生物学领域的大部分仍然在期刊发表[2]。同样现象也在中国的国家预印本平台上重现，数据表明，近年中国发表的中文期刊文章数（中国知网统计）是中国发布的预印本文章（中国科技论文在线、中国科学院预印本、中国预印本服务系统合计）的 500 多倍[3]。

随着国家预印本平台于 2022 年底落地运行，争取国家预印本平台发布文章和期刊文章一样被科研人事管理部门认可，激励大批研究者向国家预印本平台提交文章，很自然地成为国家预印本平台下一阶段建设的重要议题[4]。目前国内外预印本采用的论文质量保障机制主要有：①由预印本工作人员对研究者提交的论文进行技术性形式审查，主要包括审核作者是否在正规学术机构工作、作者履历是否具备基本的学术研究训练、软件查重以及是否有可信研究者背书。优点是审核快，缺点是并非对论文内容的评价，论文质量保障力度小。②招募志愿者[5]。志愿者机制的优点是不需要预印本平台支付报酬，缺点是学术观点分歧甚至私人矛盾可能介入其中，国际著名预印本数据库 arXiv 多次发生审核员因不同原因删掉论文而引发作者抗议事件。③读者自由点评[6]。优点是评议者可以畅所欲言，发表看法，缺点在于一些论文可能长期无人点评，另外可能出现一些作者之间互相给对方论文虚假好评，或者刻意低评与自己存在利益冲突的论文。④同行评议。由具备评议资格的评议人对论文质量进行评议。根据评议人邀请方式不同，又分为作者邀请、编辑邀请、自组织同行评议[7]。由于作者邀请评议人容易作弊或者评议人天然倾向于拔高评价，故实践中可行的是编辑邀请、自组织评议。中国国家预印本平台文章来源最多也即目前最主要子平台——中国科技论文在线——采用编辑邀请。自组织同行评议的最大特点是由算法而不是由人来完成搜索评议人、匹配评议人 - 待评材料两个同行评议的关键环节。由于算法一旦启动，不再需要人介入，故称为自组织同行评议。同行评议组织者不再具体介入搜索评议人、匹配评议人 - 待评材料，而是专注于设计与完善算法、处理同行评议过程出现的意外情况。

在评议人、待评材料相同的条件下，自组织同行评议在公正性、有效性、快捷性方面整体上优于传统的由编辑组织的同行评议[7]。作为近 10 年来同行评议模式的一种创新，自组织同行评议具有一定理论与应用价值。自组织同行评议理论上完全适用于预印本[8]，已有文献讨论将自组织同行评议应用于预印本如中国科技论文在线[9]、arXiv[10]。

自组织同行评议应用于国家预印本平台，需要解决两个核心问题：①激励研究者广泛参与问题；②关键参数设计问题。

激励研究者广泛参与问题产生的原因是同行评议普遍没有报酬，属于研究者对学术社区的义务劳动。由于期刊得到科研人事管理部门认可，因此给期刊担任文章评议人对研究者而言具有一定声誉认可价值，有助于个人职业发展，因此相对多数研究者愿意为期刊免费评议文章。预印本文章目前不被科研人事管理部门认可，国家预印本平台引入同行评议，则激励研究者参与同行评议势必成为难题。对此可引入自组织同行评议解决。自组织同行评议引入国家预印本平台

的基本思路是平台通过算法匹配一个学科领域下提交文章的研究者，让研究者之间双盲的情况下互相评议，对于连续拒评或评议超时的研究者进行处罚，处罚期间被处罚者提交和参与署名的所有论文暂停被访问、下载和被评议，同时向这些论文的其他作者通报处罚情况，通过研究者个人声誉受损与同侪压力激励评议人按时完成评议任务[8]。除通过上述处罚方式激励研究者参与外，石进等还讨论了自组织同行评议的积分激励机制[11]、付江阳等人提出自组织同行评议的通证激励[12]。

关键参数设计问题产生的原因是影响自组织同行评议性能的 4 个关键参数包括评议人休息时间、评议人拒审率、作者人数、评议资格比例，4 个关键参数取值组合不同，则自组织同行评议性能表现不同。那么，在国家预印本平台管理方要求性能下，自组织同行评议上述关键参数如何优化设计才能达到要求？

为此，本文使用论文评议完成率（CR）、评议人任务分配均衡度（TBD）、论文平均评议速度（A_r_time）3 个性能指标评价自组织同行评议表现，具体定义详见表 1，采用 Box-Behnken 实验方案，对自组织同行评议关键参数的不同取值进行试验设计，仿真得到关键参数典型组合对应的自组织同行评议在上述 3 个性能指标的表现值，再拟合得到性能指标与关键参数之间的量化方程式，国家预印本平台管理方可以根据需要设定关键参数取值，用于国家预印本平台的管理决策参考。

2 数值实验研究设计

2.1 自组织同行评议过程规则设定

为提高小同行评议概率，讨论在二级学科层面上进行。论文数与论文作者关系服从洛特卡分布，通过调整洛特卡分布参数取值从而仿真不同学科。一个仿真周期为 365 天，评议时间超出仿真周期的文章视为未评议。

自组织同行评议过程规则仿真设定：①通过调整评议人占全部作者比例来保证评议人具备必要的专业资质[10]。②3 名评议人为一组评议一篇论文（目前国内外惯例）。3 名评议人评议用时最长者为论文最终评议时间。作者不自评自己参与署名的论文。③一名评议人任务数为其提交论文数量乘以 3，如甲、乙两人分别向平台提交 1 篇、2 篇文章，则甲评议人相应任务数量为 3 篇，乙为 6 篇。提交文章多，则评议任务多。④若有多个满足条件的评议人组合，优先选择评议任务未完成数量最多的组合。⑤若存在多个评议任务未完成数量相同的评议人组合要求选择，则随机选择一个评议人组合。⑥如没有满足条件的组合，等待下一轮匹配。⑦对评议超期、多次拒评的评议人处罚为将被处罚评议人提交的待评论文直接后移到论文 - 评议人匹配序列的最后位置。

自组织同行评议与目前大多数期刊采用的编辑邀请同行评议模式的根本区别在于：自组织同行评议的评议人来源为预印本平台注册者，且为待评议论文匹配评议人时是根据一组算法规则，无需编辑介入。两者相同之处在于：都是根据摘要、关键词及作者主要研究方向，为待评议论文匹配研究领域相似的评议人。

2.2 自组织同行评议性能指标

自组织同行评议性能表现评价指标以及定义详见表 1。其中，评议任务饱和度含义是：以 3 名评议人为一组评议一篇论文为例，假设一名作者有评议资格，向国家预印本平台提交 2 篇文章要求评议，则为维持

表 1 自组织同行评议性能评价指标

Table 1 Performance evaluation indicators of self-organizing peer review

编号	名称	符号	定义
1	评议完成率	CR	一个仿真周期内，已经完成评议的论文数量占总待评议论文数的比例
2	平均评议时间	A_r_time	一个仿真周期内，处于完成评议状态的文章所用评议时间均值
3	任务分配均衡度	TBD	评议任务饱和度大于 80%的评议人数占所有具备评议资格的评议人数的比例

DOI：10.13998/j.cnki.issn1002-1248.23-0303

评议可持续，该作者需要在平台评议其他论文的评议任务量为 2×3=6。如在一个仿真周期内，此人实际完成评议论文 3 篇，则其任务饱和度为 3÷6=50%，表明平台向该评议人分配任务次数不够，或者该研究者拒审次数过多，都属于应该避免的情况。显然，任务饱和度越接近 100%越好。

2.3 参数取值

仿真参数除了 4 个关键参数（作者人数、评议资格比例、评议人休息时间、评议人拒审率）外，还包括其他一些辅助参数，全部参数具体取值如表 2 所示。

部分参数取值依据较长，表 2 补充说明如下。

（1）评议人休息时间在[15,25]区间取值依据是：据 Publons 网站完成的《全球同行评议报告》显示，论文评议时间中位数约为 14.7 天（取整为 15 天）[13]。取 5 天为步长是考虑一周工作时间为 5 天。考虑 2 月通常为 28 天，故评议人休息时间上限取值不超过 28 天。考虑步长为 5 天，为保障实验步长一致，设置休息时间上限为 25 天。

（2）据中国科学技术信息研究所 2022 年 5 月份发布的《2020 年中国科技论文统计分析》，国内发表科研论文最多的 10 个学科共计发表 28.6 万篇论文，则平均一个学科一年发表约 2~3 万篇论文。根据洛特卡定律，以论文第一作者身份一年发布一篇文章的占比最大。目前多数文章中署名作者数量不超过 3 个，综上推断得出大多数学科研究者人数约在 2~8 万人区间内。

（3）设定洛特卡参数组合时，约束条件为“撰写不同篇数文章的作者所占百分比之和等于 1”。对于绝大多数作者来说，以第一作者身份一年内撰写有一定质量的论文，通常篇数不会超过 4 篇。根据洛特卡分布撰写 1 篇、2 篇、3 篇、4 篇文章作者比例的方法如下：给定洛特卡参数 n，由广义洛特卡分布公式[14]。

$$x^{n}\cdot f(x)=C \qquad \text{公式（1）}$$

C 为参数，$f(x)$为撰写 x 篇论文的比例（x=1、2、3、4）得：

$$\sum_{i=1}^{4} f(x_i)=\sum_{i=1}^{4}\frac{C}{x_i^{\,n}}=1 \qquad \text{公式（2）}$$

根据上述归一化约束条件算出参数 C 后，可求出撰写 1 篇、2 篇、3 篇、4 篇文章作者的占比。一个洛特卡参数代表一个学科领域。

（4）一名评议人完成评议的时间上限设为 60 天的原因是：现实中多数期刊都要求评议人两个月内完成审稿。评议时间下限设为 10 天的原因是：国际及国内某些顶尖期刊要求评议人审稿时间控制在两周左右，如国内物理学顶刊《物理学报》官网“审稿政策”栏目中承诺首轮审稿通常在两周内完成。

表 2 仿真参数的取值设定

Table 2 Value setting of simulation parameters

编号	参数名称	取值范围	取值依据
1	评议人休息时间/天	[15,25]（步长：5）	具体说明详见表 2 后文
2	作者人数/人	20 000、40 000、80 000	具体说明详见表 2 后文
3	洛特卡定律参数	2~4（步长：0.25）	尽量覆盖较广学科领域范围，提高仿真研究结果的普适性
4	论文总数	由作者总数和洛特卡定律参数计算得出	
5	一年作者提交最大论文数/篇	4	为保证论文质量与评议速度，作者可以提交多篇论文，但平台一年最多评议 4 篇，多出论文自动推迟到下一年度评议
6	论文月度上传量	每月在 8%左右波动，幅度为±2%	假定投稿比较均匀，一年 12 个月，每月占比 1/12，即 8%左右
7	一名评议人完成评议时间/天	[10,60]（步长：1）	评议完成时间大致服从正态分布，使用 10~60 天的区间内正态分布随机赋值
8	评议资格比例	[50%,80%]（步长：15%）	经验设定
9	拒审率	[10%,20%]（步长：5%）	经验设定
10	一篇文章评议时间		取值为 3 名评议人评议用时最长者

2.4 实验设计

响应面分析通过实验设计取得因变量、自变量代表性数据，用多项式回归得到系统自变量和响应特征的量化关系[15]。响应面分析常用的非全因子实验设计方法有中心复合实验设计（CCD）和 Box-Behnken 实验设计（BBD）。BBD 法一个主要优点是实验参数取值不会出现超出设定范围的情况，故此处选择 BBD 方法[16]。

采用 Design-expert 软件进行设计。作者一年中提交论文数、评议时间分布、文章评议时间参数随机设定，总论文数根据洛特卡定律参数计算得出，不纳入 BBD 实验抽样。一个洛特卡定律参数代表一个学科领域，洛特卡定律参数全部讨论，不抽样。选取休息时间（A）、拒审率（B）、研究者人数（C）、评议资格比例（D）为实验因素，各自取 3 个水平，使用−1、0、+1 进行编码，分别代表低、中、高 3 种水平。

每个洛特卡定律参数均单独设计一张 BBD 表，对表内各组合进行 3 次重复实验。使用 R 语言编程，不同洛特卡定律参数 BBD 表内组合运行 50 次，按照 15 次、15 次、20 次划分为三大组。使用 R 语言中 summary() 函数和 mean()函数，计算各性能指标均值。

3 结果与讨论

3.1 实验结果

篇幅原因，本文给出洛特卡定律参数值为 2、运行 20 次后各变量取均值结果，详见表 3。其余组合结果类似。

由表 3 可知：①3 个性能指标中，CR 随因素变化的波动幅度最大，而 TBD 和 A_r_time 指标波动幅度很小，进一步考虑篇幅限制，下文只对 CR 性能指标进行响应面分析；②各实验组合 TBD 性能指标值均大于 95%，说明自组织同行评议在评议人评议任务分配方面表现良好。A_r_time 性能指标值均落在[33-35]天区间，相比较于目前期刊同行评议普遍需要 2~3 个月时间，自组织同行评议速度快很多。

3.2 CR 性能指标与 4 个关键参数的量化关系

根据表 3 数据，以 CR（y_{1j}，j=1、2、3、4、5、6、7、8、9）为因变量，以 4 个关键参数——休息时间（x_1）、拒审率（x_2）、研究者人数（x_3）、评议资格比例（x_4）——为自变量，进行函数拟合，确定量化关系。自变量数据使用无量纲化伸缩法中的归一化方法进行无量纲处理后，得到不同洛特卡定律参数取值下，4 个关键参数与 CR 的量化关系详见表 4。

在表 5 中，各模型失拟项的 Pr 值均大于 0.05，不显著，模型可信。各模型决定系数 R^2 及校正决定系数 Adj R^2 均大于 0.99，表明各模型实验实际测量值与预测值的吻合度较高。各模型 Adj R^2 和预测决定系数 Pre R^2 间高度接近。精密度 Adeq Precision 远大于 4，说明各模型拟合情况良好。各模型变异系数 C.V 小于 0.5%，表明实验结果信度高。

对于共线性问题，Design-expert 软件计算得到各洛特卡定律参数值回归方程的 VIF 值均小于 1.414，远小于通常设定的 VIF 判断值，因此不存在共线性问题。

表 4 中列出模型显示，评议资格比例变量系数绝对值是其余 3 个自变量系数绝对值的数倍，表明评议资格比例对 CR 指标的影响数倍于其他 3 个自变量。其次为休息时间，为负相关。研究者人数和拒审率对 CR 影响不显著。因此，对自组织同行评议 CR 性能指标最重要影响因素是评议资格比例，其次是休息时间，最后是研究者人数和拒审率。

以洛特卡参数值为 2 时的回归方程举例分析，保持其他 3 个参数值不变，评议资格比例数值加 1，CR 值在十位数位置变动；休息时间数值加 1，CR 值在个位数位置变动；研究者人数和拒审率数值加 1，仅变动 CR 值的百分位数，变动幅度小，可忽略不计。故在 CR 值变动时，评议资格比例贡献度最大，其次为休息时间。

3.3 关键参数交互作用

由于拒审率、研究者人数两个参数很难由国家预

DOI：10.13998/j.cnki.issn1002-1248.23-0303

表 3　洛特卡定律参数值为 2 的 BBD 实验结果

Table 3 BBD experiment results with Lotka's law parameter value of 2

编号	因素编码				自组织同行评议性能指标		
	A（休息天数/天）	B（拒审率/%）	C（研究者人数/人）	D（评议资格比例/%）	CR/%	TBD/%	A_r_time/天
1	15	10	40 000	65	63.27	98.59	34.82
2	25	10	40 000	65	61.00	95.28	33.58
3	15	20	40 000	65	63.42	98.43	33.24
4	25	20	40 000	65	61.12	95.37	33.71
5	20	15	20 000	50	47.60	96.70	33.74
6	20	15	80 000	50	47.81	96.99	33.12
7	20	15	20 000	80	77.34	97.10	34.54
8	20	15	80 000	80	77.05	97.52	34.48
9	15	15	40 000	50	48.10	98.00	33.84
10	25	15	40 000	50	46.58	95.24	34.53
11	15	15	40 000	80	78.36	98.71	34.74
12	25	15	40 000	80	75.61	95.44	34.90
13	20	10	20 000	65	62.32	96.94	33.45
14	20	20	20 000	65	62.10	96.98	34.16
15	20	10	80 000	65	62.46	97.27	34.64
16	20	20	80 000	65	62.46	97.09	33.51
17	15	15	20 000	65	63.20	98.16	34.80
18	25	15	20 000	65	60.79	95.32	33.96
19	15	15	80 000	65	63.28	98.13	34.36
20	25	15	80 000	65	61.16	95.23	33.76
21	20	10	40 000	50	47.95	97.18	33.67
22	20	20	40 000	50	47.52	96.94	33.86
23	20	10	40 000	80	77.18	97.30	34.76
24	20	20	40 000	80	76.88	97.08	34.53
25	20	15	40 000	65	62.22	96.78	33.99
26	20	15	40 000	65	62.35	97.12	33.96
27	20	15	40 000	65	62.54	97.15	33.97
28	20	15	40 000	65	62.18	97.06	33.97
29	20	15	40 000	65	62.48	97.10	33.96

印本平台管理机构调控，故此处仅讨论休息时间、评议资格比例两个参数的交互作用。

响应曲面和等高线图可直观描述试验设计中各参数间的交互作用，通常响应面的陡峭程度或等高线的形状可以在一定程度上反映出交互效应的强弱。使用 Design Expert 软件绘出响应曲面图和等高线图。篇幅原因，此处仅列出洛特卡参数值为 2 时的响应曲面图和等高线图。

根据图 1，随着休息时间减少和评议资格比例变大，CR 值随之变大，且休息时间、评议资格比例交互作用响应面图坡度较陡且等高线呈微弯曲状态，说明休息时间减少和评议资格比例变大对 CR 的影响较为显著。

表 4　不同洛特卡定律参数值下 CR 值与 4 个关键参数的量化关系

Table 4 Relationship between CR value and four key parameters under different Lotka's law parameter values

洛特卡值	CR 值与 4 个关键参数的回归方程
2	$y_{11}=62.38-1.1x_1-0.046x_2+0.072x_3+14.71x_4-0.003x_1x_2+0.066x_1x_3-0.31x_1x_4+0.053x_2x_3+0.033x_2x_4-0.12x_3x_4-0.21x_1^2+0.012x_2^2-0.027x_3^2-0.047x_4^2$
2.25	$y_{12}=62.49-1.1x_1-0.047x_2+0.071x_3+14.71x_4-0.003x_1x_2+0.065x_1x_3-0.31x_1x_4+0.052x_2x_3+0.036x_2x_4-0.1x_3x_4-0.11x_1^2+0.1x_2^2+0.069x_3^2+0.14x_4^2$
2.5	$y_{13}=62.92-0.74x_1-0.001x_2+0.04x_3+14.76x_4+0.01x_1x_2+0.044x_1x_3-0.088x_1x_4-0.06x_2x_3+0.037x_2x_4+0.098x_3x_4-023x_1^2+0.039x_2^2-0.12x_3^2-0.013x_4^2$
2.75	$y_{14}=63.21-0.74x_1-0.001x_2+0.044x_3+14.75x_4+0.009x_1x_2+0.045x_1x_3-0.087x_1x_4-0.06x_2x_3+0.041x_2x_4+0.092x_3x_4-0.22x_1^2+0.045x_2^2-0.11x_3^2-0.007x_4^2$
3	$y_{15}=63.79-0.56x_1-0.031x_2-0.047x_3+14.71x_4-0.18x_1x_2+0.063x_1x_3-0.008x_1x_4-0.078x_2x_3+0.026x_2x_4+0.048x_3x_4-0.35x_1^2-0.17x_2^2-0.15x_3^2-0.16x_4^2$
3.25	$y_{16}=63.74-0.56x_1-0.03x_2-0.048x_3+14.71x_4-0.18x_1x_2+0.063x_1x_3-0.008x_1x_4+0.078x_2x_3+0.026x_2x_4+0.048x_3x_4-0.32x_1^2-0.15x_2^2-0.13x_3^2-0.13x_4^2$
3.5	$y_{17}=63.78-0.52x_1-0.036x_2-0.084x_3+14.71x_4-0.24x_1x_2+0.021x_1x_3-0.042x_1x_4-0.079x_2x_3+0.051x_2x_4+0.046x_3x_4-0.27x_1^2-0.18x_2^2-0.12x_3^2-0.18x_4^2$
3.75	$y_{18}=64.07-0.51x_1-0.069x_2-0.047x_3+14.66x_4-0.059x_1x_2+0.047x_1x_3-0042x_1x_4+0.089x_2x_3+0.026x_2x_4-0.089x_3x_4-0.32x_1^2-0.12x_2^2-0.11x_3^2-0.14x_4^2$
4	$y_{19}=64.09-0.57x_1-0.007x_2-0.052x_3+14.75x_4-0.21x_1x_2+0.066x_1x_3+0.092x_1x_4+0.071x_2x_3+0.026x_2x_4+0.055x_3x_4-0.38x_1^2-0.16x_2^2-0.14x_3^2-0.19x_4^2$

表 5　不同洛特卡定律参数值下回归方程误差统计

Table 5 Error statistics of regression equations under different Lotka's law parameter values

洛特卡值	失拟项	R^2	Adj R^2	Pred R^2	Adeq Precision	C.V/%
2	0.480 2	0.999 8	0.999 7	0.999 3	260.629	0.27
2.25	0.206 2	0.999 9	0.999 7	0.999 3	279.800	0.25
2.5	0.415 1	0.999 9	0.999 8	0.999 6	334.659	0.20
2.75	0.604 2	0.999 9	0.999 8	0.999 6	316.227	0.22
3	0.617 9	0.999 7	0.999 5	0.999 0	196.707	0.34
3.25	0.657 5	0.999 7	0.999 5	0.998 9	193.838	0.35
3.5	0.597 4	0.999 7	0.999 5	0.998 9	193.904	0.35
3.75	0.614 9	0.999 7	0.999 5	0.998 9	196.288	0.35
4	0.411 9	0.999 7	0.999 3	0.998 3	166.873	0.41

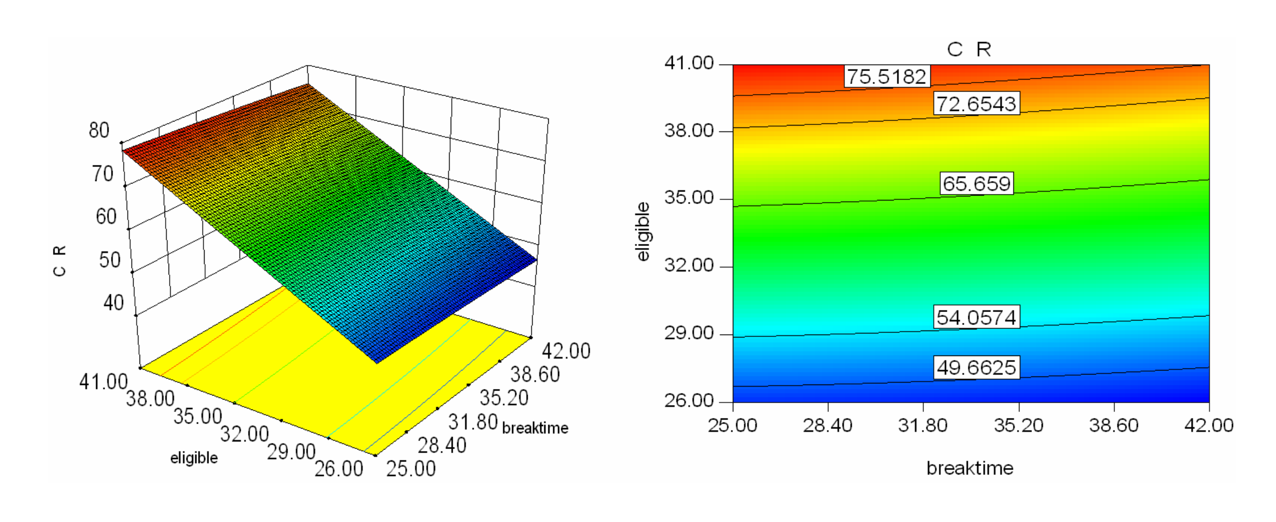


图 1　休息时间、评议资格比例对 CR 性能的交互影响

Fig.1 Interaction effects of rest time and review qualification ratio on CR performance

DOI：10.13998/j.cnki.issn1002-1248.23-0303

4 仿真方程应用：自组织同行评议关键参数优化设计示例

前文表明，休息时间、评议资格比例两个关键参数对自组织同行评议 CR 性能指标表现影响显著，同时又是国家预印本平台管理机构完全可控参数。相比之下，拒审率、研究者人数两个参数国家预印本平台管理机构不能控制。因此，管理实践中，国家预印本平台管理机构主要在一个学科提交论文作者数、经验拒审率或者运行一段时间后取得真实拒审率数据基础上，调整休息时间、评议资格比例两个参数，进行优化设计。

以某学科的作者 - 论文数分布符合洛特卡定律参数值为 2 时为例（其余洛特卡定律参数取值情况类似，不再赘述），应用表 3 中洛特卡参数值为 2 时的量化方程，如公式（3）所示。

$$y_{11}=62.38-1.1x_1-0.046x_2+0.072x_3+14.71x_4-0.003x_1x_2+0.066x_1x_3-0.31x_1x_4+0.053x_2x_3+0.033x_2x_4-0.12x_3x_4-0.21x_1^2+0.012x_2^2-0.027x_3^2-0.047x_4^2 \quad \text{公式（3）}$$

若国家预印本平台管理方预期自组织同行评议性能表现 CR 值达到 75%及以上，4 个关键参数取值可以有多种组合，篇幅原因，此处列出根据公式 3 计算的 3 种不同人数区间组合供参考：①如果作者人数落入区间[0,20000]，拒审率落入区间[10%,15%]，则其余两个关键参数取值可以确定为休息时间 20 天、评议资格比例 80%，此时 CR 值落入区间[77.24%,77.32%]；②如果作者人数落入区间[20000,40000]，拒审率落入区间[15%,20%]，则其余两个关键参数取值可以确定为休息时间 15 天、评议资格比例 80%，此时 CR 值落入区间[78.32%,78.34%]；③如果作者人数落入区间[40000,80000]，拒审率落入区间[15%,20%]，则其余两个关键参数取值可以确定为休息时间 25 天、评议资格比例 80%，此时 CR 值落入区间[75.51%,75.56%]。其余情况，根据文中所得量化方程再进行计算即可。

5 结　语

本研究主要结论：①评议资格比例、评议人休息时间对自组织同行评议性能影响较大，特别是评议资格比例设定重要性远高于其他 3 个关键参数；②量化方程中，作者人数变量系数绝对值非常小，表明作者人数变化对自组织评议性能影响极低，近似于可忽略。

本文对国家预印本平台建设问题的贡献为：通过仿真为国家预印本平台采用自组织同行评议保证论文质量提供了数值实验论据，得到的仿真方程可为国家预印本平台管理机构决策提供量化支持。需要强调的是，由于本文作者掌握算力资源有限，故仿真精度还有提高空间，如需要且具备算力资源的情况下，可以按照本文方法和参数进行更加精细的仿真计算。

参考文献：

[1] 李广建. 对国家科研论文和科技信息高端交流平台的理解及研究建议[J]. 图书与情报, 2021(6): 1-4.
LI G J. Understanding and research suggestions on the high-end exchange platform of national scientific research papers and scientific and technological information[J]. Library & information, 2021(6): 1-4.
[2] 孙异凡, 陈一, 蒋子可, 等. 开放科学视域下预印本认可政策研究[J]. 数字图书馆论坛, 2021(6): 2-12.
SUN Y F, CHEN Y, JIANG Z K, et al. Research on the recognition policy of preprints from the perspective of open science[J]. Digital library forum, 2021(6): 2-12.
[3] 王凌峰, 程灵慧. 预印本与期刊: 互补合作到优势融合[J/OL]. 情报杂志: 1-7[2023-06-19]. http://kns.cnki.net/kcms/detail/61.1167.G3.20230714.1106.006.html.
WANG L F, CHENG L H. Preprint and journal: Complementary cooperation to advantage integration[J/OL]. Journal of intelligence: 1-7[2023-08-19].http://kns.cnki.net/kcms/detail/61.1167.G3.20230714.1106.006.html.
[4] 唐碧群, 王凌峰. 预印本 2.0 应用于国家科研论文高端交流平台建设的设想[J]. 图书情报导刊, 2021, 6(9): 20-24.
TANG B Q, WANG L F. A proposal of preprint 2.0 applied to national high-level scientific paper platform construction[J]. Journal of library and information science, 2021, 6(9): 20-24.
[5] 宋永辉, 马廷灿, 刘静羽. 志愿者参与下国际预印本平台学术质量控制方法调研与启示——以 arXiv 与 RePEc 平台为例[J]. 中国

科技期刊研究, 2023, 34(2): 119–126.

SONG Y H, MA T C, LIU J Y. Roles of volunteers in academic quality control of international preprint platforms: Taking arXiv and RePEc as examples[J]. Chinese journal of scientific and technical periodicals, 2023, 34(2): 119–126.

[6] 祝忠明. 国家高端交流平台建设体系及其开放科学实践策略[J]. 图书与情报, 2021(6): 21–30.

ZHU Z M. Open science platforms for the development of high–end communication platforms in China[J]. Library & information, 2021(6): 21–30.

[7] 唐碧群, 王凌峰. 自组织同行评议: 技术赋能同行评议的产生、设计与多场景应用[J]. 图书情报导刊, 2022, 7(11): 63–68.

TANG B Q, WANG L F. Self–organized peer review: The generation, design and multi–scenario applications of technology–enpowered peer review[J]. Journal of library and information science, 2022, 7(11): 63–68.

[8] 王凌峰, 孙英潮. E–prints 预印本数据库的自组织同行评议模式设计[J]. 现代情报, 2016, 36(5): 9–15.

WANG L F, SUN Y C. Design of self–organizing peer review mechanism under E–prints database environment[J]. Journal of modern information, 2016, 36(5): 9–15.

[9] 唐碧群, 杨硕, 王凌峰. 自组织同行评议在预印本数据库中的应用——以“中国科技论文在线”为例[J]. 图书情报导刊, 2018, 3(4): 26–31.

TANG B Q, YANG S, WANG L F. Application of self–organizing peer review in preprint database: Taking "Chinese science papers online" as an example[J]. Journal of library and information science, 2018, 3(4): 26–31.

[10] WANG L F, ZHAN Y Q. A conceptual peer review model for arXiv and other preprint databases [J]. Learned publishing, 2019(3): 213–219.

[11] 石进, 苗杰, 李明. 面向预印本系统的自组织同行评议及激励机制研究[J]. 现代情报, 2019, 39(12): 88–100.

SHI J, MIAO J, LI M. Research on self–organization peer review and incentive mechanism for pre–printed system[J]. Journal of modern information, 2019, 39(12): 88–100.

[12] 付江阳, 贺颖. 预印本自组织同行评议的通证激励机制研究[J]. 中国科技期刊研究, 2021, 32(1): 23–27.

FU J Y, HE Y. Token incentive mechanism of self–organizing peer review in preprint[J]. Chinese journal of scientific and technical periodicals, 2021, 32(1): 23–27.

[13] PUBLONS. Publons' global state of peer review 2018[R]. Global State of peer review report – Clarivate, London, UK: Publons, 2018.

[14] 张贤澳, 李美文. 广义洛特卡定律参数特征的研究[J]. 情报探索, 1996(2): 10–15.

ZHANG X A, LI M W. Study on parameter characteristics of generalized lotca's law[J]. Information research, 1996(2): 10–15.

[15] 郭勤涛, 张令弥, 费庆国. 用于确定性计算仿真的响应面法及其试验设计研究[J]. 航空学报, 2006, 27(1): 55–61.

GUO Q T, ZHANG L M, FEI Q G. Response surface method and its experimental design for deterministic computer simulation[J]. Acta aeronautica et astronautica Sinica, 2006, 27(1): 55–61.

[16] 王永菲, 王成国. 响应面法的理论与应用[J]. 中央民族大学学报(自然科学版), 2005, 14(3): 236–240.

WANG Y F, WANG C G. The application of response surface methodology[J]. Journal of the central university for nationalities (natural science edition), 2005, 14(3): 236–240.

研究论文
DOI：10.13998/j.cnki.issn1002-1248.23-0303

Key Parameter Optimization Design of Self-organizing Peer Review in National Preprint Publishing Platform Based on Response Surface Analysis

WNAG Lingfeng, WANG Shenpeng
(Business School, Guilin University of Electronic Technology, Guilin 541004)

Abstract: [Purpose/Significance] The National Preprint Publishing Platform of China had been put into operation by the end of 2022. In order to ensure the quality of papers posted on the Platform and encourage a large number of researchers to submit articles to the Platform, it is necessary to introduce a paper quality assurance mechanism into the Platform. There are deficiencies in existing quality assurance mechanism for a preprint publishing platform. Self-organizing peer review has gradually attracted attention in recent years. In order to achieve the performance expected by the management agency of the National Preprint Publishing Platform, it is necessary to study the optimization design of key parameters of self-organizing peer review. [Method/Process] Design-expert software has been used for Box Behnken experimental design. The number of papers submitted by authors within a year, the distribution of review time, and the parameters of article review time are randomly set. The total number of papers is calculated according to the parameters of Lotka's law, and is not included in the Box-Behnken experimental sampling. We selected four key parameters of rest time, rejection rate, number of researchers, and review qualification ratio as independent variables, took three levels for each variable, and used -1, 0, +1 for coding, representing three levels of low, medium, and high, respectively. The three dependent variables for self-organizing peer review performance evaluation are the completion rate of paper review (CR), the balance of reviewer task allocation (TBD), and the average review time of the paper (A_r_time). A Box-Behnken experiment table was designed separately for each Lotka's law parameter, three repeated experiments were conducted on each key parameter combination in the table, and the summary() function and mean() function in R language were used to calculate the average performance indicators of each key parameter combination. Then, the response surface analysis was carried out on the Box-Behnken sampling data, and the quantitative relationship expression between the dependent variable of performance evaluation and the independent variable of key parameters was obtained. [Results/Conclusions] The ratio of review qualifications and reviewer rest time have greater impact on the performance of self-organized peer review, especially the ratio of review qualifications is much more important than other three key parameters. By combining different values of key parameter independent variables, the National Preprint Publishing Platform can achieve the expected self-organized peer review performance, ensuring the average quality of papers published by the Platform. In addition, because of our limited knowledge of computational resources, there is still room for improvement in simulation accuracy. If equipped with more powerful computational resources, more precise simulation calculation can be carried out based on the methods and parameters proposed in this paper.

Keywords: National Pre-print Publishing Platform; paper quality; self-organizing peer review; box-behnken design; response surface analysis